A distortion of very–high–redshift galaxy number counts by gravitational lensing

J. Stuart B. Wyithe[1], Haojing Yan[2], Rogier A. Windhorst[3], Shude Mao[4,5]

[1]School of Physics, University of Melbourne, Parkville, Victoria 3010, Australia

[2]Center for Cosmology and AstroParticle Physics, The Ohio State University, Columbus, OH 43210

[3]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404

[4]Jodrell Bank Centre for Astrophysics, University of Manchester, Manchester, M13 9PL, UK

[5]National Astronomical Observatories of China, Chinese Academy of Sciences, Beijing 100012, China

The observed number counts of high-redshift galaxy candidates[1−8] have been used to build up a statistical description of star-forming activity at redshift $z \gtrsim 7$, when galaxies reionized theUniverse[1,2,9,10]. Standard models[11] predict that a high incidence of gravitational lensing will probably distort measurements of flux and number of these earliest galaxies. The raw probability of this happening has been estimated to be ~ 0.5 per cent (refs 11, 12), but can be larger owing to observational biases. Here we report that gravitational lensing is likely to dominate the observed properties of galaxies with redshifts of $z \gtrsim 12$, when the instrumental limiting magnitude is expected to be brighter than the characteristic magnitude of the galaxy sample. The number counts could be modified by an order of magnitude, with most galaxies being part of multiply imaged systems, located less than 1 arcsec from brighter foreground galaxies at $z \approx 2$. This lens-induced

association of high-redshift and foreground galaxies has perhaps already been observed among a sample of galaxy candidates identified at $z \approx 10.6$. Future surveys will need to be designed to account for a significant gravitational lensing bias in high-redshift galaxy samples.

Along random lines-of-sight, the *raw* probability (or optical depth) for multiple imaging of objects at high redshifts — owing to gravitational lensing by individual foreground field galaxies[11,12] — is $\simeq 0.5\%$. However, all galaxy populations are observed to have a characteristic luminosity ($L_\star$), brighter than which galaxy numbers drop exponentially, and below which numbers rise with a very steep power-law slope[1,4,6]. The potential for gravitational lensing to modify the observed statistics therefore increases dramatically, owing to the magnification of numerous, intrinsically faint galaxies to observed fluxes that are above the survey limit. This effect, which is known as *magnification bias*[13], leads to an excess of gravitationally lensed galaxies among flux-limited samples. Magnification bias is expected to be particularly significant at high redshifts ($z \gtrsim 8$), where current observations may only be probing the exponential tail of the LF[4], so that the number density could be rising very rapidly towards the detection limit. Indeed, multiply imaged candidates at $z \gtrsim 7$ have already been discovered behind foreground clusters via targeted searches[14–16], demonstrating this to be an efficient method for finding faint high redshift galaxies[14,17].

We assess magnification bias among high redshift galaxies assuming singular, spherical, isothermal gravitational lenses, which produce one or two images, and designate the apparent magnitude of the more magnified image (the only image in the absence of lensing) as $m_{\rm AB,1}$. We then calculate, as a function of the assumed characteristic luminosity (expressed in terms of absolute magnitude $M_\star$), the fraction of galaxies brighter than the magnitude limit ($m_{\rm lim}$) for the Hubble Ultra Deep Field (*HUDF*) that would be multiply imaged (designated $F_{\rm lens}$). Such curves are shown at $z = 6$, 7, 8.6 and 10.6 in panel **a** of Figure 1. The superimposed solid and open points correspond to lens fractions for different estimates[1,4] of $M_\star$ at these redshifts. At $z \simeq 6-7$, we expect only $\simeq 1$ percent of galaxies

to be lensed. At $z \simeq 8-10$, however, we expect a lensed fraction of a few to a few tens of percent, depending on the true value of $M_\star$. Note that since current survey limits are significantly fainter than $M_\star$ at $z \simeq 6-7$, the lens fraction is quite insensitive to $M_\star$. However, at higher redshifts where the survey limits might be much closer to $M_\star$, the lensing fraction is very sensitive to its uncertain value.

Predictions for a significant lens fraction at $z \gtrsim 8$ stand in apparent contrast to the fact that no image pairs have been identified in the *HUDF*. However, we find the probability that a multiply imaged galaxy, with observed $m_{\rm AB,1}$ has a corresponding *second* image with $m_{\rm AB,2} < m_{\rm lim}$ (i.e. detectable with the *HUDF* data) to be only $\simeq 10\%$, even for galaxies that are one magnitude brighter than $m_{\rm lim}$ (see Supplementary Information). Thus, as shown in panel **b** of Figure 1, the fraction of galaxies ($F_{\rm mult}$) that are detected as multiply imaged systems in the *HUDF* is an order of magnitude lower than the true lensed fraction. Although this fraction would increase somewhat if elliptical lenses were included in our analysis, multiply imaged systems are not expected to be observed in the current data. On the other hand, magnification bias also leads to a concentration of high redshift sources — both singly and multiply imaged — around foreground galaxies[18–20]. The resulting correlation between high redshift candidates and bright foreground galaxies therefore offers an alternative avenue to observing the effect of gravitational lensing. A schematic diagram illustrating this point, as well as magnification bias, is included as Supplementary Figure 1.

To quantify this correlation, we first determine the distribution of separations between random lines-of-sight and the nearest bright ($H \leq 25$ mag) foreground galaxy in the *HUDF*, measured as the angular distance to the centroid. This is shown by the dotted black line in panel **c** of Figure 1. This distribution can be compared to the predictions of our model (dashed line in panel **c** of Figure 1). If the candidate sample consists of both multiply imaged *and* unmagnified galaxies, then the observed distribution of separations should be a weighted sum of the random and the lensed line-of-sight distributions. The correct weighting is the probability for gravitational lensing, $F_{\rm lens}$. Two examples

are shown in panel **c** of Figure 1. The fraction of galaxies found within $\Delta\theta \simeq 1-2$ arcseconds of a foreground galaxy is very sensitive to the characteristic luminosity if $M_\star \gtrsim -19$ mag, providing a potential observable for the influence of lensing on the number counts of $z \gtrsim 8$ candidates.

For comparison with the lensing predictions, we have measured the distribution of separations between a sample of $z \sim 10.6$ candidates[4] and their nearest bright ($H \leq 25$ mag) foreground galaxy. Comparing the distributions, we find that these candidates are observed to be closer to bright foreground galaxies than are random lines-of-sight. On the other hand, the candidates are found at larger separations from foreground galaxies than would be predicted if they were all multiply imaged. Quantitatively, the Kolmogorov-Smirnov probabilities between the observed distributions and the *all-random* model or the *all-lensed* model (see Supplementary Information) indicate that both models are rejected at high significance. This suggests that a fraction of candidates may be gravitationally lensed. Moreover, we have generated the distribution of redshifts for foreground galaxies found within $\Delta\theta < 1.5$ arcseconds of the $z \sim 10.6$ candidates. These distributions are consistent with the distribution of gravitational lens redshifts, while the redshift distribution of all bright foreground galaxies are not, which supports the hypothesis that foreground galaxies are lensing a fraction of the $z \sim 10.6$ candidates into the observed sample.

With the introduction of the James Webb Space Telescope (*JWST*), galaxy surveys will be undertaken out to even higher redshifts, well into the epoch of First Light[21]. Panels **a** and **b** of Figure 2 show $F_{\rm lens}$ as a function of $M_\star$ out to $z = 20$. The flux limits correspond to an ultra-deep survey ($m_{\rm lim} = 31.4$ mag), and a medium-deep survey ($m_{\rm lim} = 29.4$ mag). The evolution of the characteristic luminosity is unknown at these unexplored redshifts. For comparison, we therefore plot squares corresponding to estimates of $M_\star$ based on an extrapolation from lower redshift *HUDF* data[1]. Figure 2 shows that in ultra-deep *JWST* surveys for First Light objects at $z \gtrsim 14$, more than $F_{\rm lens} \sim 10\%$ of the candidates could be lensed. In much shallower *JWST* surveys that only sample the exponential tail of the Schechter LF, a lensed object fraction of $F_{\rm lens} \sim 10\%$ could be seen at

redshifts as low as $z \sim 8$–10. However at $z \gtrsim 14$, the lensed fraction in such surveys could be much higher, and may even represent the majority of observed galaxies. Surveys with *JWST* will therefore need to be carefully planned and analyzed to account for the influence of foreground lensing galaxies.

As in the case of the *HUDF* the fraction of galaxies that will be detected as multiply imaged systems by *JWST* is significantly lower than the true multiple image fraction. However, as the multiple image fraction becomes very large at high redshifts, observed doubles could become common; larger than $F_{\rm mult} \sim 10\%$ at redshifts $z \gtrsim 12$ in a medium-deep ($m_{\rm AB} < 29.4$ mag) *JWST* survey, and $z \gtrsim 16$ in an ultra-deep ($m_{\rm AB} < 31.4$ mag) survey. Panels **e** and **f** present the predicted distributions of separation for galaxies discovered by *JWST* from bright foreground galaxies. If the observed evolution in $M_\star$ continues to higher redshift, then the spatial distribution of high redshift galaxies relative to foreground galaxies will depart from random at redshifts $z \gtrsim 14$ for ultra-deep surveys, and at $z \gtrsim 10$ for medium-deep surveys with *JWST*. A crucial prediction is that the majority of very high redshift galaxies discovered with *JWST* may be located less than 1 arc-second from a bright foreground galaxy, and will have been gravitationally magnified into the sample.

A key goal for *JWST* will be to measure the number counts of high redshift candidates, and to construct luminosity functions (LF) in order to build up a statistical description of star-forming activity in galaxies. LFs describing the density of sources per unit luminosity are parametrised by a Schechter function[22], $\Psi(L) \propto (L/L_\star)^\alpha \exp{(-L/L_\star)} 1/L_\star$, including free parameters for the power-law slope at low luminosities (α), and the characteristic absolute AB-magnitude [$M_{\rm AB} - M_\star = -2.5 \log_{10}(L/L_\star)$] brighter than which galaxy numbers drop exponentially. Importantly, gravitational lensing has the potential to significantly modify the observed LF from its intrinsic shape[23]. In particular, at very high luminosities in the exponential tail of the Schechter function, the LF shape can be modified from exponential to power-law, since gravitational lensing magnifies numerous faint sources to apparently higher luminosities. Figure 3 shows that the shapes of LFs near the flux limit are not affected by gravitational lensing at $z \approx 6 - 8$. However, *if* the evolution of the galaxy LF

continues into the reionisation era (we assume an extrapolation of the fitting formulae based on candidates discovered in and around the *HUDF*[1]), then we find that *JWST* will measure LFs that are significantly modified by lensing at redshifts above $z \sim 14$ and $z \sim 10$ in its ultra-deep and medium-deep surveys, respectively.

Our results imply that while published LF's at $z \gtrsim 7$ are not currently corrected for a potential gravitational lensing bias, such corrections will need to be prescribed in detail for future surveys that aim to measure the build-up of stellar mass among the first galaxies using *JWST*. In particular, studies of the high redshift LF will require good understanding of the magnification bias for high redshift galaxies, in order to correct for gravitational lensing and uncover its true unlensed shape at $z \gtrsim 12$. Of particular importance will be the unknown evolution of $M_\star$, which could be influenced (for example) by supernovae feedback from population-III stars[24], in addition to hierarchical clustering and formation. Gravitational lensing could magnify $z \gtrsim 10$–12 objects to flux levels that will allow spectroscopic observations using JWST and the largest ground-based near-IR spectrographs. A further implication of our analysis is that gravitational lensing could be used to probe the shape of the high redshift LF at luminosities that are not otherwise accessible[12], using the association of high redshift galaxy candidates and foreground galaxies, combined with careful modelling of the gravitational lensing bias.

Acknowledgments:
The authors thank K.-H. Chae. JSBW was supported in part by a QE-II fellowship and grants from the Australian Research Council. HY acknowledges supports of the long-term fellowship program of the Center for Cosmology and AstroParticle Physics (CCAPP) at The Ohio State University. HY and RAW were supported by grants from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc. RAW was supported by a NASA JWST Interdisciplinary Scientist grant.

References

[1] Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Labbe, I., Trenti, M., van Dokkum, P., Franx, M., Stiavelli, M., Carollo, C. M., Magee, D., and Gonzalez, V. UV Luminosity Functions from 113 $z \sim 7$ and $z \sim 8$ Lyman-Break Galaxies in the ultra-deep HUDF09 and wide-area ERS WFC3/IR Observations. *ArXiv e-prints* **1006.4360**, (2010).

[2] Lorenzoni, S., Bunker, A., Wilkins, S., Stanway, E., Jarvis, M., and Caruana, J. Candidate $z \sim 8-9$ Galaxies from WFC3 Imaging. *ArXiv e-prints* **1006.3545**, (2010).

[3] Bouwens, R. J., Illingworth, G. D., Labbe, I., Oesch, P. A., Carollo, M., Trenti, M., van Dokkum, P. G., Franx, M., Stiavelli, M., Gonzalez, V., and Magee, D. Constraints on the First Galaxies: $z \sim 10$ Galaxy Candidates from HST WFC3/IR. *ArXiv e-prints* **0912.4263**, (2009).

[4] Yan, H., Windhorst, R., Hathi, N., Cohen, S., Ryan, R., O'Connell, R., and McCarthy, P. Galaxy Formation In The Reionization Epoch As Hinted By Wide Field Camera 3 Observations Of The Hubble Ultra Deep Field. *Res. in Astron. & Astropys.* **10**, 867–904 (2010).

[5] Bouwens, R. J., Illingworth, G. D., Thompson, R. I., Blakeslee, J. P., Dickinson, M. E., Broadhurst, T. J., Eisenstein, D. J., Fan, X., Franx, M., Meurer, G., and van Dokkum, P. Star Formation at $z \sim 6$: The Hubble Ultra Deep Parallel Fields. *Astrophys. J. Lett.* **606**, L25–L28, (2004).

[6] Yan, H. and Windhorst, R. A. Candidates of $z \sim 5.5-7$ Galaxies in the Hubble Space Telescope Ultra Deep Field. *Astrophys. J. Lett.* **612**, L93–L96, (2004).

[7] Wilkins, S. M., Bunker, A. J., Ellis, R. S., Stark, D., Stanway, E. R., Chiu, K., Lorenzoni, S., and Jarvis, M. J. Probing L_* Lyman-break galaxies at $z \sim 7$ in GOODS-South with WFC3 on Hubble Space Telescope. *Mon. Not. R. Astron. Soc* **403**, 938–944, (2010).

[8] McLure, R. J., Dunlop, J. S., Cirasuolo, M., Koekemoer, A. M., Sabbi, E., Stark, D. P., Targett, T. A., and Ellis, R. S. Galaxies at $z = 6-9$ from the WFC3/IR imaging of the Hubble Ultra Deep Field. *Mon. Not. R. Astron. Soc* **403**, 960–983, (2010).

[9] Yan, H. and Windhorst, R. A. The Major Sources of the Cosmic Reionizing Background at $z \sim 6$. *Astrophys. J. Lett.* **600**, L1–L5, (2004).

[10] Trenti, M., Stiavelli, M., Bouwens, R. J., Oesch, P., Shull, J. M., Illingworth, G. D., Bradley, L. D., and Carollo, C. M. The Galaxy Luminosity Function During the Reionization Epoch. *Astrophys. J. Lett.* **714**, L202–L207, (2010).

[11] Barkana, R. and Loeb, A. High-Redshift Galaxies: Their Predicted Size and Surface Brightness Distributions and Their Gravitational Lensing Probability. *Astrophys. J.* **531**, 613–623, (2000).

[12] Comerford, J. M., Haiman, Z., and Schaye, J. Constraining the Redshift $z \sim 6$ Quasar Luminosity Function Using Gravitational Lensing. *Astrophys. J.* **580**, 63–72, (2002).

[13] Turner, E. L., Ostriker, J. P., and Gott, III, J. R. The statistics of gravitational lenses - The distributions of image angular separations and lens redshifts. *Astrophys. J.* **284**, 1–22, (1984).

[14] Richard, J., Stark, D. P., Ellis, R. S., George, M. R., Egami, E., Kneib, J., and Smith, G. P. A Hubble and Spitzer Space Telescope Survey for Gravitationally Lensed Galaxies: Further Evidence for a Significant Population of Low-Luminosity Galaxies beyond $z = 7$. *Astrophys. J.* **685**, 705–724, (2008).

[15] Bradley, L. D., Bouwens, R. J., Ford, H. C., Illingworth, G. D., Jee, M. J., Benítez, N., Broadhurst, T. J., Franx, M., Frye, B. L., Infante, L., Motta, V., Rosati, P., White, R. L., and Zheng, W. Discovery of a Very Bright Strongly Lensed Galaxy Candidate at $z \sim 7.6$. *Astrophys. J.* **678**, 647–654, (2008).

[16] Zheng, W., Bradley, L. D., Bouwens, R. J., Ford, H. C., Illingworth, G. D., Benítez, N., Broadhurst, T., Frye, B., Infante, L., Jee, M. J., Motta, V., Shu, X. W., and Zitrin, A. Bright Strongly Lensed Galaxies at Redshift $z \sim 6-7$ behind the Clusters Abell 1703 and CL0024+16. *Astrophys. J.* **697**, 1907–1917, (2009).

[17] Bouwens, R. J., Illingworth, G. D., Bradley, L. D., Ford, H., Franx, M., Zheng, W., Broadhurst, T., Coe, D., and Jee, M. J. $z \sim 7-10$ Galaxies Behind Lensing Clusters: Contrast with Field Search Results. *Astrophys. J.* **690**, 1764–1771, (2009).

[18] Webster, R. L., Hewett, P. C., Harding, M. E., and Wegner, G. A. Detection of statistical gravitational lensing by foreground mass distributions. *Nature* **336**, 358–360, (1988).

[19] Nollenberg, J. G. and Williams, L. L. R. Galaxy-Quasar Correlations between APM Galaxies and Hamburg-ESO QSOs. *Astrophys. J.* **634**, 793–805, (2005).

[20] Scranton, R., Ménard, B., Richards, G. T., Nichol, R. C., Myers, A. D., Jain, B., Gray, A., Bartelmann, M., Brunner, R. J., Connolly, A. J., Gunn, J. E., Sheth, R. K., Bahcall, N. A., Brinkman, J., Loveday, J., Schneider, D. P., Thakar, A., and York, D. G. Detection of Cosmic Magnification with the Sloan Digital Sky Survey. *Astrophys. J.* **633**, 589–602, (2005).

[21] Windhorst, R. A., Cohen, S. H., Jansen, R. A., Conselice, C., and Yan, H. How JWST can measure first light, reionization and galaxy assembly. *New Astron. Review* **50**, 113–120, (2006).

[22] Schechter, P. An analytic expression for the luminosity function for galaxies. *Astrophys. J.* **203**, 297–306, (1976).

[23] Pei, Y. C. Magnification of Quasars by Cosmologically Distributed Gravitational Lenses. *Astrophys. J.* **440**, 485–500, (1995).

[24] Greif, T. H., Johnson, J. L., Bromm, V., and Klessen, R. S. The First Supernova Explosions: Energetics, Feedback, and Chemical Enrichment. *Astrophys. J.* **670**, 1–14, (2007).

[25] Hathi, N. P., Ryan, Jr., R. E., Cohen, S. H., Yan, H., Windhorst, R. A., McCarthy, P. J., O'Connell, R. W., Koekemoer, A. M., Rutkowski, M. J., Balick, B., Bond, H. E., Calzetti, D., Disney, M. J., Dopita, M. A., Frogel, J. A., Hall, D. N. B., Holtzman, J. A., Kimble, R. A., Paresce, F., Saha, A., Silk, J. I., Trauger, J. T., Walker, A. R., Whitmore, B. C., and Young, E. T. UV-dropout Galaxies in the GOODS-South Field from WFC3 Early Release Science Observations. *Astrophys. J.* **720**, 1708–1716, (2010).

[26] Oguri, M., Inada, N., Strauss, M. A., Kochanek, C. S., Richards, G. T., Schneider, D. P., Becker, R. H., Fukugita, M., Gregg, M. D., Hall, P. B., Hennawi, J. F., Johnston, D. E., Kayo, I., Keeton, C. R., Pindor, B., Shin, M., Turner, E. L., White, R. L., York, D. G., Anderson, S. F., Bahcall, N. A., Brunner, R. J., Burles, S., Castander, F. J., Chiu, K., Clocchiatti, A., Eisenstein, D., Frieman, J. A., Kawano, Y., Lupton, R., Morokuma, T., Rix, H., Scranton, R., and Sheldon, E. S. The Sloan Digital Sky Survey Quasar Lens Search. III. Constraints on Dark Energy from the Third Data Release Quasar Lens Catalog. *Astron. J.* **135**, 512–519, (2008).

[27] Choi, Y., Park, C., and Vogeley, M. S. Internal and Collective Properties of Galaxies in the Sloan Digital Sky Survey. *Astrophys. J.* **658**, 884–897, (2007).

[28] Huterer, D., Keeton, C. R., and Ma, C. Effects of Ellipticity and Shear on Gravitational Lens Statistics. *Astrophys. J.* **624**, 34–45, (2005).

[29] Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Limon, M., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Meyer, S. S., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. *Astrophys. J. Sup.* **180**, 330–376, (2009).

[30] Windhorst, R. A., Hathi, N. P., Cohen, S. H., Jansen, R. A., Kawata, D., Driver, S. P., and Gibson, B. High resolution science with high redshift galaxies. *Advan. Space Res.* **41**, 1965–1971 (2008).

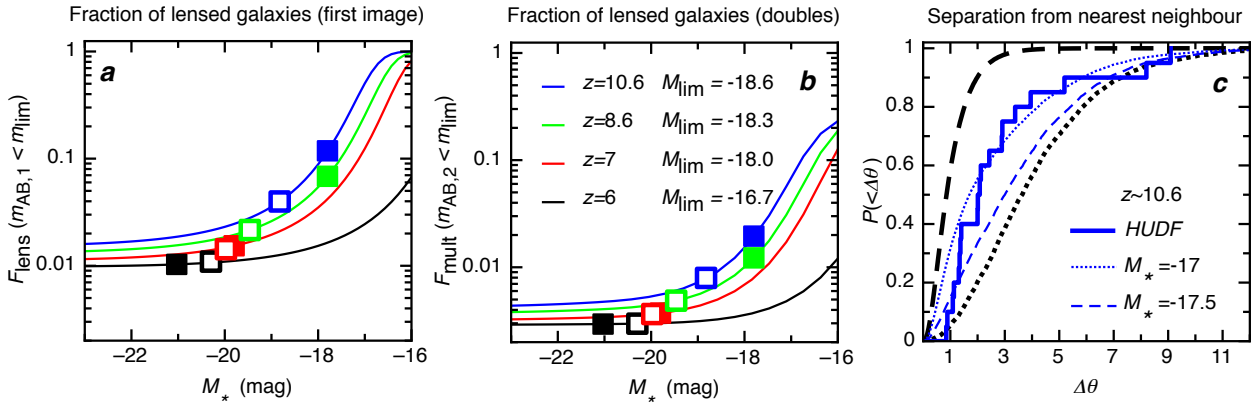


Figure 1. **Gravitational lens fractions among candidate high redshift *HUDF* galaxies. Panel a:** The fraction of multiply-imaged high redshift galaxies. **Panel b:** The fraction of high redshift galaxies in which multiple images could be detected in the *HUDF*. **Panel c:** The probability distribution of image separations (at $z \simeq 10.6$) relative to the nearest bright foreground galaxy, in the cases of random lines-of-sight (black dotted line), of gravitational lenses (black dashed line), and for composite distributions computed for two (faint) values of $M_\star$. Also shown is the distribution of measured separations for twenty $z \sim 10.6$ candidates[4] in the *HUDF* (stepped blue histogram). Lyman-break galaxy candidates have been selected with median redshifts of $z \simeq 6$, $z \simeq 7$, $z \simeq 8.6$ and $z \simeq 10.6$. At $z \simeq 6$, candidate selection using the Advanced Camera for Surveys reaches[6] $m_{\rm lim} = 30$ mag (absolute magnitude $M_{\rm lim} = -16.7$ mag). At higher redshifts, objects in the *WFC3 HUDF* data can be selected[4] to $m_{\rm lim} \simeq 29.0$ mag, corresponding to $M_{\rm lim} = -18.0$, -18.3 and -18.6 mag at $z \simeq 7$, 8.6 and 10.6. The open squares correspond to lens fractions given the fitting formula[1,25] $M_\star \simeq -21 + 0.32 \times (z - 3.8)$. The solid squares represent alternative estimates[4,6] of $M_\star$. The model for gravitational lensing[26] is based on the velocity dispersion function of galaxies[27]. Galaxy mass distributions are modelled as Singular Isothermal Spheres, and we assume a constant co-moving density of lenses. Elliptical lenses would not significantly alter the cross-section[28], but would provide additional images, and so increase the fraction of observed galaxies that are lensed. We assume a Schechter LF[22], with power-law slope[1] $\alpha = -2$. A change of 0.3 in α leads to a 40% change in the lens probability. We have used the cosmology based on 7-year results from the *WMAP* satellite[29] throughout this *Letter*.

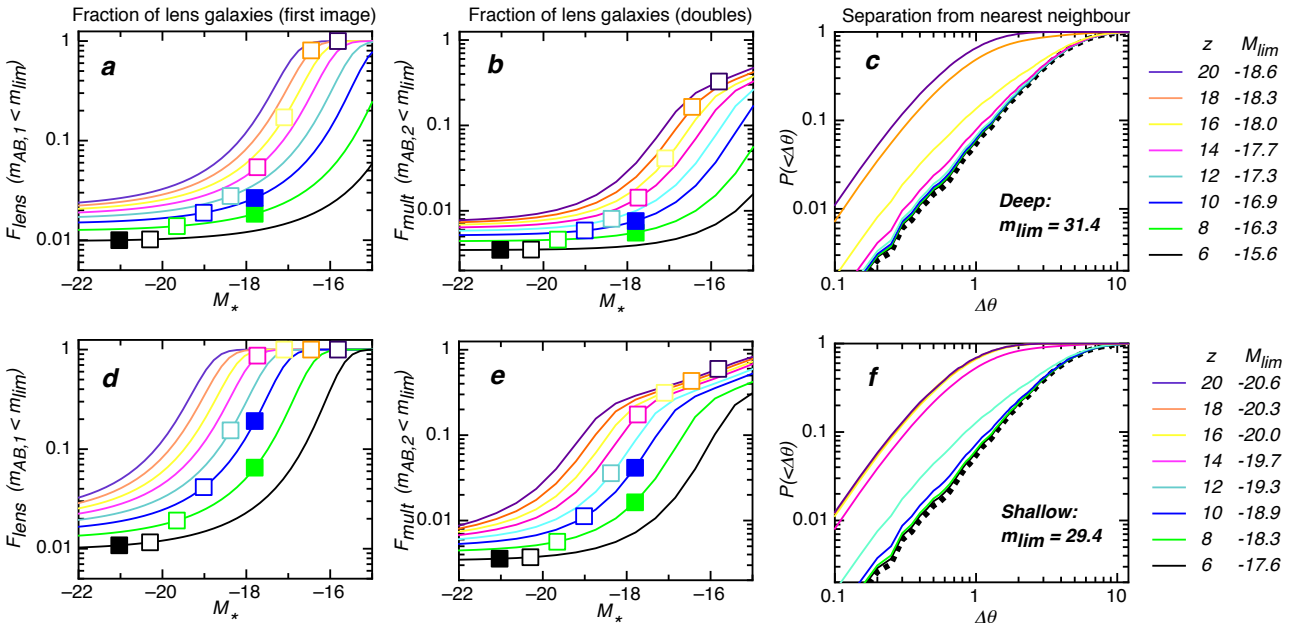


Figure 2. **Probabilities for multiple imaging of high redshift galaxies to be observed with *JWST*.** The panels mirror those of Figure 1, but with examples of limiting magnitudes and redshifts appropriate for both an *ultra-deep* survey ($m_{\rm AB} < 31.4$ mag, ~ 1 nJy), and a *medium-deep* survey ($m_{\rm AB} < 29.4$ mag) with *JWST*. The corresponding limiting absolute magnitudes are listed. **Panels a-b:** The fraction of observed galaxies that have multiple images. The superimposed solid and open points correspond to lens fractions given a faint value[4] of $M_\star = -17.8$ at $z \sim 8.6$ and $z \sim 10.6$, and a fitting formula $M_\star(z)$ based on lower redshift data, respectively[1,25]. The latter is extrapolated to high redshift where data does not yet exist. **Panels c-d:** The fraction of high redshift galaxies in which multiple images could be detected by *JWST*. **Panels e-f:** The probability distribution of image separations relative to the nearest bright foreground galaxy, in the cases of random lines-of-sight (black dotted line), and for composite distributions computed for values of $M_\star$ extrapolated from observations in the *HUDF* using the previously mentioned fitting formula[1,25]. We note that imaging surveys with *JWST* will be working at the diffraction limit (~ 0.08 arcseconds resolution FWHM) at $\sim 2\,\mu$m. This resolution is higher than is currently available in the *HUDF* near-IR images, where candidates have been selected in close proximity to bright foreground galaxies, and hence high redshift candidates will also be detectable close to foreground galaxies.

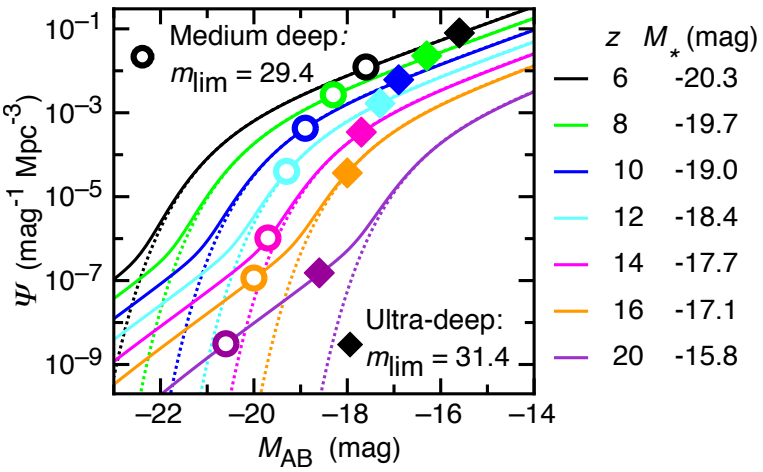


Figure 3. **Gravitational lens induced modification of the bright end of the high redshift galaxy luminosity function to be observed with *JWST*.** Thin curves present the intrinsic LF (Ψ), and solid curves the observed LF following modification from gravitational lensing. For simplicity, a uniform magnification was assumed outside regions of sky that are multiply-imaged, with a value such that flux is conserved over the whole sky. The parameters describing the LF are extrapolated to high redshift, where data does not yet exist, assuming fitting formulae based on data from the *HUDF*[1,25]. Of particular relevance are the values of $M_\star$, which are listed. The solid and open points show the luminosities and densities of the faintest galaxies to be observed with *JWST*, assuming limiting magnitudes appropriate for both an *ultra-deep JWST* survey ($m_{\rm AB} < 31.4$ mag), and a *medium-deep JWST* survey ($m_{\rm AB} < 29.4$ mag). The probability for gravitational lensing will become of order unity in the steep exponential parts of the LF at sufficiently high redshifts. This *gravitational forest* should not to be confused with the purely mathematical effects of image crowding that makes the detection and de-blending of faint objects harder at progressively fainter fluxes[30]. These latter effects are referred to as either the *instrumental confusion limit* — when the instrumental resolution is not good enough to statistically distinguish *all* faint background objects from brighter foreground objects — or the *natural confusion limit* — when the instrumental resolution is good enough to distinguish faint background objects from brighter foreground objects, but the images are so deep that objects start overlapping because of their own intrinsic sizes. The *HUDF* and *JWST* images are in the latter regime[30], and as argued in this *Letter*, likely have the *additional* fundamental limitation that *gravitational lensing* will magnify a non-negligible fraction of faint objects into the sample.

—

SUPPLEMENTARY INFORMATION

—

The *Letter to Nature* estimates the probability for gravitational lensing among high redshift galaxies, with emphasis on current surveys using the *HUDF*, and future surveys to be carried out using *JWST*. The *Letter* also uses the observed distributions of separation between very high redshift galaxy candidates in the *HUDF* and foreground galaxies, to show that a significant fraction of these objects are likely to be gravitationally lensed. The following sections expand the brief descriptions of the modelling and interpretation that can be found in the *Letter to Nature*.

1 Schematic picture of magnification bias and foreground galaxy correlation

In Supplementary Figure 1, we present a schematic representation of a portion of the *HUDF*, which shows how magnification bias leads to a correlation between foreground galaxies and high redshift candidates. Panel **a** shows a representation of the Schechter function[22], which describes the luminosity function (LF) of high redshift galaxies. The limiting absolute magnitude $M_{\rm lim}$ and characteristic magnitude $M_\star$ are shown for reference. Gravitational lensing magnifies sources relative to their intrinsic luminosity, and draws intrinsically faint galaxies into the flux limited sample. Since faint galaxies are much more common than bright galaxies, the number of sources per unit area in regions of lensing magnification is significantly higher. This leads to a bias of sources near foreground galaxies. To illustrate this effect on the high redshift galaxy samples in the *HUDF*, we sketch in panel **b** a portion of the sky approximately 10 arcseconds across. In this panel, background sources (i.e. high redshift galaxies) are shown in red and foreground galaxies (those near $z \simeq 1-2$) in blue. The faint galaxies (with $M_{AB} > M_{\rm lim}$) are signified by open symbols, while the closed symbols signify bright galaxies with $M_{AB} < M_{\rm lim}$. The black dotted disks denote regions of sky where background sources will be multiply-imaged by a foreground galaxy. For illustration, this schematic representation overestimates the total lensing cross-section, which is $\simeq 0.5\%$, by a factor of $\simeq 10$. The typical angular scale of these regions is 1 arc-second. We show those faint galaxies that lie within these lensing regions in green. In panel **c,** the faint galaxies that are close enough to bright foreground galaxies to be multiply-imaged (shown in green), producing in general a bright image with $M_{AB} < M_{\rm lim}$, and an undetected faint image with $M_{AB} > M_{\rm lim}$. Finally, the observed association of high redshift galaxies with bright foreground galaxies — once gravitational lensing bias has been accounted for — is shown in panel **d**. In this example 2 of the 5 observed high redshift galaxies ($M_{AB} < M_{\rm lim}$) have entered the sample owing to gravitational magnification, and have close alignment with foreground galaxies as a result. In this case we find 40% of high redshift galaxies within $\simeq 1$ arc-second of bright foreground galaxies, even though the observed density of bright foreground objects is 1 per 20 square arcseconds. We note that gravitational lensing can also lower the observed density of sources on the sky that have neighbouring foreground galaxies by magnifying the angular extent of the image plane relative to the source plane. This effect, which is usually referred to as *depletion*, is not dominant when the LF is steep, as is the case for high redshift galaxies.

2 Lens model

We refer to the a-priori probability for a galaxy at redshift $z_{\rm gal}$ to be multiply-imaged by an intervening foreground galaxy as the multiple image optical depth[17]

$$\tau_{\rm m} = \int_0^{z_{\rm gal}} \frac{d\tau_{\rm m}}{dz} dz, \tag{1}$$

where

$$\frac{d\tau_{\rm m}}{dz} = \int d\sigma \Phi(\sigma, z)(1+z)^3 \frac{cdt}{dz} \pi D_{\rm d}^2 \theta_{\rm ER}^2(\sigma, z), \tag{2}$$

$\theta_{\rm ER}$ is the Einstein radius as a function of velocity dispersion σ and redshift z, $D_{\rm d}$ is the angular diameter distance to the lens, and t is time. To calculate $\tau_{\rm m}$, we use the expression for the angular Einstein radius for a Singular Isothermal Sphere (SIS)

$$\theta_{\rm ER}(\sigma, z) = 0.9'' \ \frac{D_{\rm ds}}{D_{\rm s}} \left(\frac{\sigma}{161 {\rm km/s}} \right)^2, \tag{3}$$

where $D_{\rm s}$ and $D_{\rm ds}$ are the angular diameter distances to the source, and between the lens and source, respectively.

To evaluate $\Phi(\sigma, z)$, we first assume[26] the *Sloan Digital Sky Survey* (SDSS) velocity dispersion function[27] $\Phi_{\rm SDSS}(\sigma)$

$$\Phi_{\rm SDSS}(\sigma) d\sigma = \Phi_\star \left(\frac{\sigma}{\sigma_\star} \right) \frac{\exp\left[-(\sigma/\sigma_\star)^\beta\right]}{\Gamma(\alpha/\beta)} \beta \frac{d\sigma}{\sigma}, \tag{4}$$

where $\Phi_\star = 2 \times 10^{-3} {\rm Mpc}^{-3}$, $\alpha = 2.32$, $\beta = 2.67$ and $\sigma_\star = 161\,{\rm km/s}$. We further assume that the lens population has a constant co-moving density $\Phi(\sigma, z) = \Phi_{\rm SDSS}(\sigma)$. Although the density of galaxies must decline at high redshift, this approximation is reasonable, since most lensing occurs at $z \lesssim 1.5$. The uncertainty in predictions of the lens fraction owing to the unknown evolution of the velocity dispersion function is approximately a factor of two[26,31] . We note that this prescription gives a lensing cross-section for $z \simeq 2$ quasars that is consistent with the SDSS analysis[26], which is an observational requirement. The lens model assumes that galaxy velocity dispersions reach down to as low as $\sigma = 10$km/s. However, as the lensing neighbours are selected by velocity dispersion, the distribution of lensed separations is not sensitive to the assumed cutoff, because the lens cross-section is proportional to velocity dispersion to the fourth power (σ^4).

We have utilised a simple lens model. In particular we have not included non-spherical lens distributions, which produce four rather than two image lenses in some cases. Indeed, empirical estimates for the fraction of quasar lenses that have four images of about 40% have been obtained from the homogeneous CLASS sample (`http://www.aoc.nrao.edu/~smyers/class.html`), and of about 15% from the Sloan Digital Sky Survey[32] . While the predicted four-image to two-image ratio depends on the ellipticity of the lensing galaxies[32] , the ellipticity is found not to significantly influence the overall cross-section for multiple imaging[28,33,34]. On the other hand, the magnification bias can be larger for an elliptical lens, which would increase the expected multiple imaging rate[28]. Moreover, the additional images in a four image lens would increase the fraction of observed candidates that are part of a multiply-imaged galaxy. Using spherical lenses for our estimates is therefore conservative with respect the expected influence of gravitational lensing on samples of high redshift galaxy candidates, both in terms of the number of lenses predicted and the association between high redshift candidates and bright foreground galaxies. We note here that the lens population for $z \sim 8-10$ candidates is at higher redshift than the lens galaxies responsible for the aforementioned samples. However the measured ellipticity distribution is nearly constant over a very wide range of flux and redshift[35] . Thus, we argue that since our simple model provides a good statistical description of the available data, neglecting elliptical lenses is reasonable, particularly given the range of other uncertainties.

2.1 magnification bias

Flux limited samples are subject to magnification bias, which increases the relative probability that detected galaxies are gravitationally lensed[17], and concentrates sources in a flux limited sample around foreground objects[18]. Yan et al.[4] have observed a number of $z \simeq 8-10$ candidates that have neighbouring bright foreground galaxies. As discussed in the *Letter*, this correlation is likely to be the manifestation of these effects. The magnification bias for sources with observed luminosities between L and $L + dL$ is

$$B(L) = \frac{\int_{\mu_{\rm min}}^{\mu_{\rm max}} \frac{d\mu}{\mu} \frac{dP}{d\mu} \Psi(L/\mu)}{\Psi(L)}, \tag{5}$$

while the corresponding overall magnification bias in a flux limited sample is

$$B_{\rm lens} = \frac{\int_{\mu_{\rm min}}^{\mu_{\rm max}} d\mu \int_{L_{\rm lim}}^{\infty} dL \frac{dP}{d\mu} \Psi(L/\mu)}{\int_{L_{\rm lim}}^{\infty} dL \frac{dP}{d\mu} \Psi(L)}, \tag{6}$$

where $dP/d\mu$ is the probability distribution for magnification (μ) within the range $\mu_{\rm min} < \mu < \mu_{\rm max}$. Of relevance for high redshift surveys in the *HUDF* or with *JWST* (which have an angular resolution much better than the image separation[30]) is the magnification distribution for the brighter image

$$\frac{dP_{\rm m,1}}{d\mu} = \frac{2}{(\mu-1)^3} \quad \text{for} \quad 2 < \mu < \infty. \tag{7}$$

We adopt a Schechter[22] function for the LF

$$\Psi(L)dL = \Psi_\star \left(\frac{L}{L_\star}\right)^\alpha \exp\left(-\frac{L}{L_\star}\right)\frac{dL}{L_\star}, \tag{8}$$

where $\Psi_\star$ is the characteristic density in Mpc^{-3}, and α is the power-law slope at luminosities below the characteristic break at $L_\star$. Below, and in the *Letter*, we quote the characteristic luminosity in terms of the absolute magnitude $M_\star = M + 2.5 \log_{10} L/L_\star$.

2.2 gravitationally lensed luminosity function

We note that in the presence of significant gravitational lensing, the LF can be modified from its intrinsic form[23], leading to a power-law slope at the bright-end of -3 (as shown in Figure 3 of the *Letter*). The modified LF can be estimated by modelling the overall magnification distribution using the probability distribution for magnification of multiply-imaged sources over a fraction $\tau_{\rm m}$ of the sky, combined with a de-magnification $\mu_{\rm demag} = (1 - \langle\mu_{\rm mult}\rangle\tau_{\rm m})/(1-\tau_{\rm m})$ elsewhere. Here $\langle\mu_{\rm mult}\rangle = 4$ is the mean magnification of multiply-imaged sources, and $\mu_{\rm demag}$ has been calculated in order to conserve flux on the cosmic sphere centred on an observer. The modified LF can then be approximated using the expression

$$\Psi_{\rm obs}(L) = (1-\tau_{\rm m})\frac{1}{\mu_{\rm demag}}\Psi(L/\mu_{\rm demag}) + \tau_{\rm m}\int_0^\infty d\mu \frac{1}{\mu}\left(\frac{dP_{\rm m,1}}{d\mu} + \frac{dP_{\rm m,2}}{d\mu}\right)\Psi(L/\mu), \tag{9}$$

where $dP_{\rm m,2}/d\mu = 2/(\mu+1)^3$ for $0 < \mu < \infty$, is the probability distribution for the second image. We approximate the true magnification distribution by using a constant value of $\mu_{\rm demag}$ in regions of no multiple imaging. This is valid for the modification of the LF at luminosities much brighter than $M_\star$, in which we are interested in this work.

3 Lensing predictions for high redshift surveys

We summarise the predictions of our lensing model in Supplementary Figure 2. As shown in panel **a**, the lensing optical depth rises toward high redshift[11], and is 4-5 times as large for sources at $z \simeq 6$ as at $z \simeq 1.5$. It doubles again from $z = 6$ to $z = 20$, so that at $z > 10$ the multiple imaging fraction is greater than 0.5%, even in the absence of magnification bias. Panel **b** shows the magnification bias as a function of the difference between $M_\star$ and the survey limit in absolute magnitude $M_{\rm lim}$. At low redshifts, deep surveys can probe well below $M_\star$, so that the magnification bias is dominated by the power-law slope (α) of the Schechter function at low luminosities, and the resulting bias is of order unity. At very high redshifts, however, current surveys can only reach $M_\star$ or even brighter, and hence the bias can be much higher (tens or hundreds) owing to the exponential nature of the LF sampled. We next combine the optical depth $\tau_{\rm m}$ with the bias $B_{\rm lens,1}$ to find the multiple image fraction $F_{\rm lens} = B_{\rm lens,1}\tau_{\rm m}/(B_{\rm lens,1}\tau_{\rm m} + (1 - \tau_{\rm m}))$, where we have assumed the bias of those galaxies which are not multiply-imaged to be unity. In panel **c** we plot contours of $F_{\rm lens}$ as a function of z and $(M_\star - M_{\rm lim})$. Surveys at low redshift ($z \lesssim 3$), with limits fainter than $M_\star$, should have multiple image fractions below 1%. However, at higher redshifts the lens fraction can be much higher. For example, a survey at $z \gtrsim 6$ that reaches only 1 magnitude brighter than $M_\star$ could have a lens fraction of 10%. Current and future surveys at $z > 6$ with *HST* and *JWST* lie in this upper-right portion of panel **c**. Only ultradeep surveys with *JWST* that reach well below $M_\star$ at $z \gtrsim 10$ will have their lensing fraction drop well below 10% again.

4 High redshift galaxy candidate samples

To compare the predictions of our model with samples of high redshift galaxy candidates, we investigate samples from the *HUDF* compiled by Yan et al.[4]. These and other authors have employed the Lyman-break (or *dropout*) technique to select galaxies at $z \gtrsim 7$ in the *HUDF*. We note that the major colour criteria used to select the samples of Yan et al.[4] are very similar to those employed by other groups including Bouwens et al.[1,3]. However the overlap of individual candidates among the samples from these two teams is small. In particular, none of J-dropouts compiled by Yan et al.[4] are among the three J-dropouts presented by Bouwens et al.[1]. There could be a range of reasons for this disjoint. With respect to our current work, we note that one reason for the difference in sample selection could be the choice of whether to include candidates near bright foreground objects. By construction, the samples of Yan et al.[4] were not biased against regions around foreground objects, indicating that *if* gravitationally lensed, multiply-imaged galaxies do exist in the *HUDF* at $z \sim 8-10$, then they would be selected. We therefore concentrate here on the predicted gravitational lensing statistics for these samples.

The $z \approx 8.6$ sample used to discuss the gravitational lensing of galaxies in the *HUDF* as part of this work consists of 15 Y-dropouts (spanning the redshift range of $7.7 \lesssim z \lesssim 9.4$), while the $z \approx 10.6$ sample consists of 20 J-dropouts (spanning $9.4 \lesssim z \lesssim 11.8$). These objects are all very faint, and have magnitudes ranging from $M_{\rm AB} = 28.0 - 29.0$.

4.1 lensing predictions for $z \simeq 8 - 10$ candidates

We have calculated multiple-imaging probabilities for the $z \simeq 8 - 10$ samples[4] as a function of galaxy absolute magnitude assuming $M_\star = -17.8$ mag. These results can be used to discuss lensing probabilities for individual $z \simeq 8 - 10$ dropout candidates[4] in more detail.

Panel **a** of Supplementary Figure 3 shows the probability that a galaxy with absolute magnitude $M_{AB,1}$ is multiply-imaged. At $z \simeq 6-7$, only galaxies much brighter than $M_{AB,1} < -21$ mag have a significant chance of being lensed. However, at $z \simeq 8-10$ galaxies as faint as $M_{AB,1} \simeq -19$ mag have

a substantial lens fraction. Of course, these are just statements reflecting the relative brightness of $M_{\rm lim}$ and M_*. Our results suggest that a number of $z \simeq 8-10$ galaxies detected in the *HUDF* should be multiply-imaged. On the other hand, we note that we have not identified any image pairs in the *HUDF*. Panel **b** shows the probability that a *lensed* galaxy with observed $M_{AB,1}$ has a corresponding *second* image with $M_{AB,2} < M_{\rm lim}$, such that it is also detectable above the *HUDF* flux limit. For this probability to be large ($\gtrsim 50\%$), the detected image must be more than $\simeq 1$ mag brighter than $M_{\rm lim}$. Panel **c** shows the fraction of galaxies that are part of a lensed pair in which *both* images are detectable, $[F_{\rm dbl} = F_{\rm lens} \times P(M_{AB,2} < M_{\rm lim}|M_{AB,1})]$. We find that at $z \simeq 6-7$, only galaxies that are several magnitudes brighter than $M_{\rm lim}$ have a reasonable chance (few-10%) of being observed as a multiple image system. However, at $z \simeq 8-10$, this probability increases to $\gtrsim 10\%$ for galaxies that are only a magnitude brighter than $M_{\rm lim}$.

We note that the predicted fraction would increase if we modelled elliptical lenses which can have more than two images. We roughly estimate the fraction in this case by noting that a four-image lens typically has either two bright images of approximately equal magnification where the source is near a fold caustic, or three bright images with the central one having a magnification equal to the sum of the other two[36] . Thus, close to the detection limit, we expect either the two bright images, or only the brightest of three bright images would be detected for typical four-image lenses. We therefore argue that for the (empirically observed) 15-40% of cases where the lens has four images, the fraction of multiply-imaged systems in which more than one image is detected will increase by at most a factor of approximately two.

In Supplementary Figure 3, we have superimposed squares to show probabilities for individual galaxy candidates in the *HUDF* [4] . We use $M_\star = -17.8$ mag estimated by Yan et al.[4] as an example. By summing probabilities for individual galaxy candidates in the Yan et al.[4] sample, we calculate the (mean) expected number of lensed systems, finding $\langle N_{\rm lens}\rangle = 0.8 \pm 0.1$ and $\langle N_{\rm lens}\rangle = 1.7 \pm 0.2$ among the 15 and 20 candidates at $z \simeq 8.6$ and $z \simeq 10.6$, respectively. If the true $M_\star$ value is fainter, these numbers will be higher. A Poisson distribution with mean $\langle N_{\rm lens}\rangle = 2.5$ implies that at least one lens pair would be found among the observed $z \simeq 8-10$ sample in 92% of cases, which stands in apparent contrast to the fact that no image pairs have been identified in the *HUDF*. However, we find the probability that a *lensed* galaxy with observed $m_{AB,1}$ has a corresponding *second* image with $m_{AB,2} < m_{\rm lim}$ (i.e. detectable with the *HUDF* data) to be only $\simeq 10\%$, even for galaxies that are one magnitude brighter than $M_{\rm lim}$. Here we neglect the caveat that secondary images could fall on top of the foreground galaxies, which would further reduce the chance of their being observed. We estimate that the number of systems that would be observed as doubles (i.e. both images detected) to be $\langle N_{\rm dbl}\rangle = 0.2 \pm 0.06$ and $\langle N_{\rm dbl}\rangle = 0.4 \pm 0.1$ at $z = 8.6$ and 10.6, respectively. A Poisson distribution with mean $\langle N_{\rm dbl}\rangle = 0.6$ implies that the observed $z \simeq 8-10$ sample would not contain any doubles in most (55%) cases. Thus, with $M_\star = -17.8$, we find that $N_{\rm lens} \simeq 2-3$ of the detected galaxies in each redshift range should be multiply-imaged, but do not necessarily expect any of these to be identified as multiple image systems. On the other hand, an even fainter value of $M_\star = -17.3$ (-17.1) implies that at least one double would be observed in 90% (99%) of cases, imposing an upper-limit of $M_\star \lesssim -17$ at $z \sim 8-10$. We note that the values of $M_\star$ — as measured — could also be biased by gravitational lensing (see Figure 3 of the *Letter*). Currently, none of the published LFs at $z \gtrsim 7$ are corrected for the potential lensing bias. However it is clear from the results presented in our *Letter*, that such corrections will need to be prescribed in detail in the future.

The mean magnification of detected lensed images (with $M_\star = -17.8$) is $\langle\mu\rangle \simeq 6$, indicating that gravitational lensing in these samples would lead to over-estimates of the luminosity density at $z \simeq 8.6$ and $z \simeq 10.6$ of $\simeq 50\%$ and $\simeq 80\%$, respectively, if the magnification is neglected. Since the axis ratio of lensed images is equal to the magnification for an SIS, this also implies that the lensed images should be significantly elongated, and indeed some candidates appear to have this property[4] . However, the signal-to-noise for the detected candidates is too low to draw quantitative conclusions.

5 Distribution of lensed separations

As shown in the previous section, we find that in most cases only the more magnified image will be brighter than the detection threshold. We therefore calculate the expected distribution of angular separation between a lens galaxy and the brighter of the two images. The apparent angular separation of the bright image with magnification μ from the center of a lensing SIS at redshift $z < z_{\rm gal}$ is

$$\Delta\theta_{\rm lens}(\mu, z) = \left(1 + \frac{1}{\mu - 1}\right)\theta_{\rm ER}(z). \tag{10}$$

Using this expression we evaluate the probability distribution for the separation of bright images of image pairs from the lensing galaxy

$$\frac{dP}{d\Delta\theta} \propto \int_0^{z_{\rm gal}} dz \int_2^{\infty} d\mu \int_{L_{\rm lim}}^{\infty} dL \frac{d\tau_{\rm m}}{dz} \frac{dP_{\rm m,1}}{d\mu} \Psi(L/\mu)\delta_{\rm dir}([\Delta\theta - \Delta\theta_{\rm lens}(\mu, z)], \tag{11}$$

where $\delta_{\rm dir}$ is the Dirac delta function, and $L_{\rm lim}$ is the unlensed luminosity corresponding to the survey flux limit.

6 Observed correlation between high redshift candidates and foreground galaxies

For comparison with the lensing predictions, we measure the distribution of separations between $z \simeq 8-10$ candidates[4] and their nearest bright ($H \leq 25$ mag) foreground galaxy. The red histograms in panels **a** of Supplementary Figures 4 and 5 show the cumulative distributions of this separation for the $z \simeq 8.6$ and $z \simeq 10.6$ candidates, respectively. Comparing the distributions in these two panels with the random line-of-sight and lensed predictions, two trends are obvious. Firstly, these $z \simeq 8-10$ candidates are observed to be closer to bright foreground galaxies than are random lines-of-sight. On the other hand, the candidates are found at larger separations from foreground galaxies than would be predicted if they were all multiply-imaged. Quantitatively, the Kolmogorov-Smirnov probabilities ($P_{\rm KS}$) between the observed distributions and the *all-random* model or the *all-lensed* model (labeled in the figure) indicate that either model is rejected at high significance. This suggests that a fraction of these candidates may be gravitationally lensed.

For illustration, the thick black lines in panels **a** of Supplementary Figures 4 and 5 show the composite distributions corresponding to multiple image lens fractions of $F_{\rm lens} = 0.2$ and 0.4, at $z \simeq 8.6$ and 10.6, respectively. These provide an excellent fit to the data ($P_{\rm KS}$ values labeled in Supplementary Figures 4 and 5). Panels **b** of Supplementary Figures 4 and 5 show the differential distributions of the observed angular separations (red), as well as the corresponding composite models (thick black) and the random distributions (dotted black). The latter demonstrates that the largest observed separations can be attributed to a random distribution.

We next examine the redshift distributions of the nearest bright foreground galaxies, using spectrophotometric redshift estimates[37] . The cumulative distributions are compared in panels **c** of Supplementary Figures 4 and 5 for the $z \simeq 8.6$ and 10.6 candidates, respectively. The red histograms are the distributions for the neighbours of the high redshift candidates, the dotted black lines are distributions for the neighbours of random lines-of-sight, and the dashed black lines are distributions for the expected gravitational lens redshift[17]. The redshift distributions of the foreground galaxies associated with the full samples of $z \simeq 8-10$ candidates cannot be differentiated from those associated with random lines-of-sight. In addition, for the $z \simeq 10.6$ case in particular, foreground galaxy redshifts are found not to be drawn from a lensed galaxy population. However, the lens angular separation cuts off sharply at $\Delta\theta \simeq 1.5$ arcseconds. We therefore generate the distribution of redshifts only for foreground galaxies found within $\Delta\theta < 1.5$ arcseconds of the $z \simeq 8-10$ candidates,

which are shown as the blue histograms in these two panels. These distributions are consistent with the distribution of gravitational lens redshifts, which supports the hypothesis that many close candidate–foreground galaxy pairs in this sample result from magnification bias. In panels **c** and **d** of Supplementary Figures 4 and 5, we show the model redshift distributions corresponding to the values $F_{\rm lens} = 0.2$ and 0.4 for candidates at $z \simeq 8.6$ and 10.6, respectively (thick black lines). These again provide an excellent fit to the data, which, when taken together with the correlation between high redshift and foreground galaxy positions, provides compelling evidence for a significant lens fraction among the $z \gtrsim 8$ galaxy candidates, since these foreground galaxies were selected only on the basis of their alignment with high redshift candidates.

References

[31] Chae, K. Galaxy evolution from strong-lensing statistics: the differential evolution of the velocity dispersion function in concord with the Λ cold dark matter paradigm. *Mon. Not. R. Astron. Soc.* **402**, 2031–2048, (2010).

[32] Oguri, M. Is there a quad problem among optical gravitational lenses? *New J. of Phys.* **9**, 442–450, (2007).

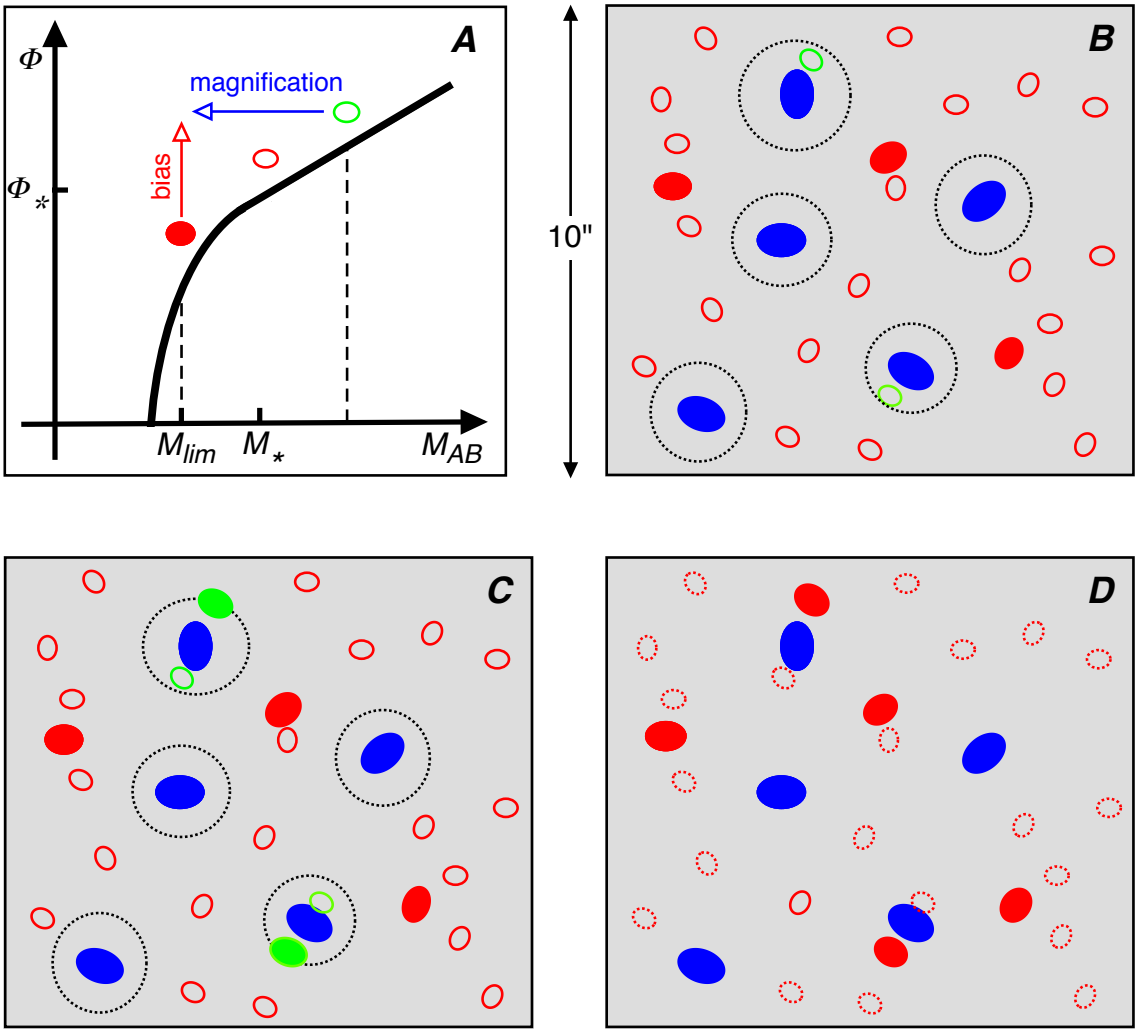
[33] Blandford, R. D. and Kochanek, C. S. Gravitational imaging by isolated elliptical potential wells. I - Cross sections. II - Probability distributions. *Astrophys. J.* **321**, 658–675, (1987).

[34] Kochanek, C. S. The flat-spectrum radio luminosity function, gravitational lensing, galaxy ellipticities, and cosmology. *Astrophys. J.* **473**, 595–609, (1996)

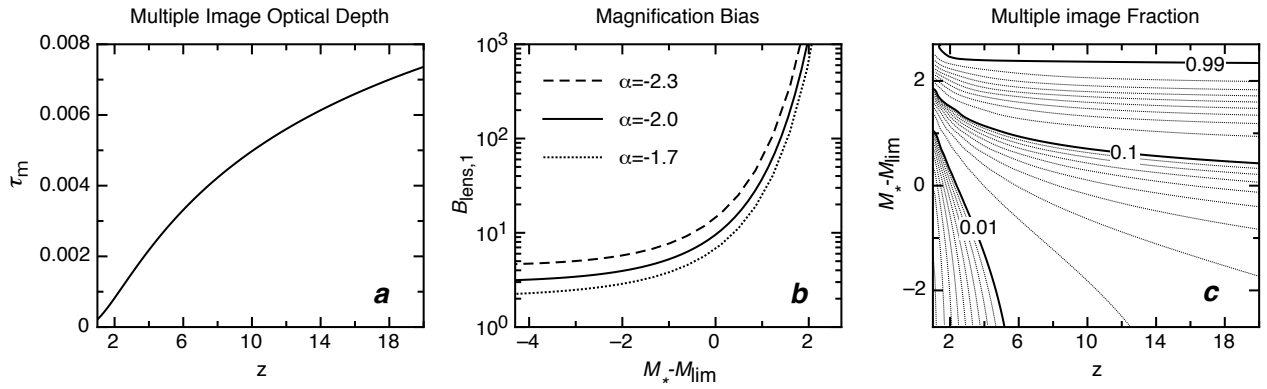
[35] Odewahn, S., Burstein, D., and Windhorst, R. A. The axis ratio distribution of local and distant galaxies. *Astron. J.***114**, 2219–2231, (1997)

[36] Mao, S. Gravitational microlensing by a single star plus external shear. *Astrophys. J.* **389**, 63–67, (1992).

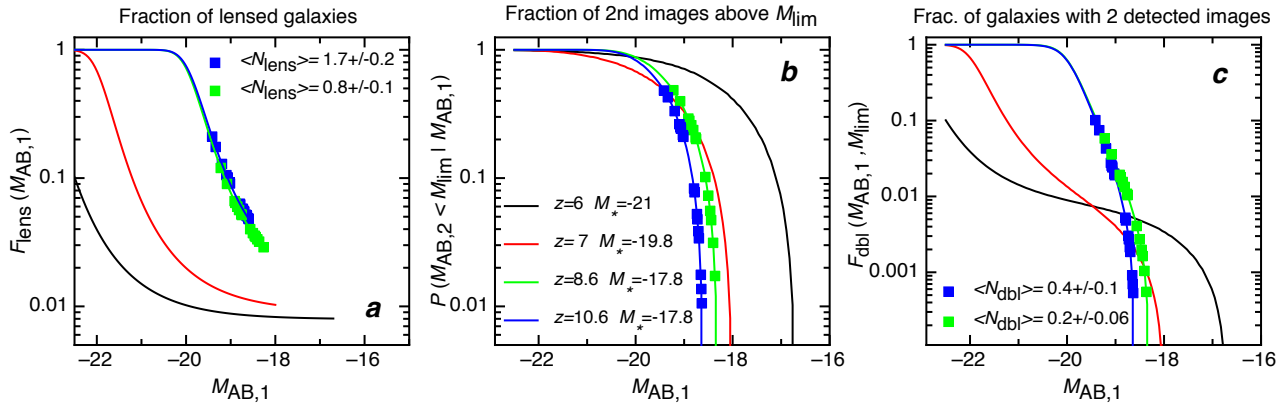
[37] Ryan, Jr., R. E., Hathi, N. P., Cohen, S. H., Malhotra, S., Rhoads, J., Windhorst, R. A., Budavári, T., Pirzkal, N., Xu, C., Panagia, N., Moustakas, L. A., di Serego Alighieri, S., and Yan, H. The Galaxy Luminosity Function at $z \sim 1$ in the *HUDF*: Probing the Dwarf Population. *Astrophys. J.* **668**, 839–845, (2007).



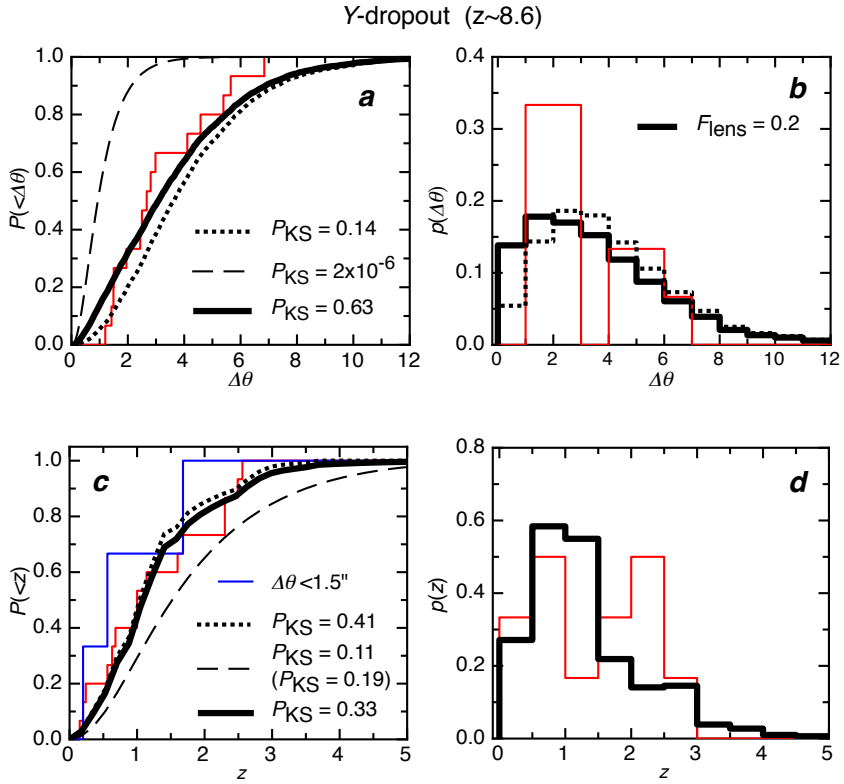
Supplementary Figure 1. **Schematic representation showing how magnification bias leads to an association between foreground galaxies and high redshift candidates. Panel a:** The Schechter LF of high redshift galaxies. **Panel b:** High redshift galaxies (red) and foreground galaxies (blue). Faint galaxies (those with $M_{AB} > M_{\rm lim}$) are signified by open symbols, while the closed symbols signify bright galaxies with $M_{AB} < M_{\rm lim}$. The black dotted disks denote regions of sky where background sources will be multiply imaged by the foreground galaxy. Faint background galaxies that lie within these lensing regions are shown in green. **Panel c:** The lensed faint galaxies are multiply-imaged, producing a bright image with $M_{AB} < M_{\rm lim}$, and an undetected faint image with $M_{AB} > M_{\rm lim}$. Galaxies located near the lines of sight to foreground galaxies that are not multiply imaged, are deflected to larger separations, resulting in a lowering of observed source density (an effect known as depletion). **Panel d:** The correlation of observed high redshift galaxies (solid red symbols) with bright foreground galaxies once gravitational lensing bias has been accounted for. The depletion effect is opposite in sign to the correlation introduced through strong lensing, but is sub-dominant in the case of high redshift galaxies.



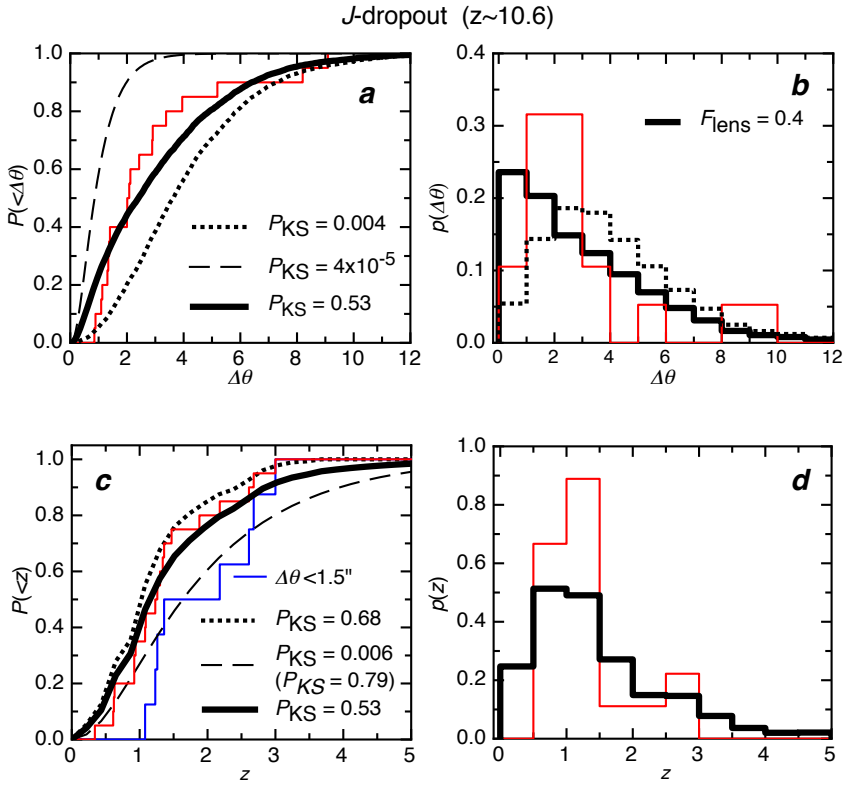
Supplementary Figure 2. **Probabilities for multiple imaging of high redshift galaxies. Panel a:** The lensing optical depth as a function of redshift. **Panel b:** The magnification bias as a function of the difference between $M_\star$ and the limiting survey absolute magnitude $M_{\rm lim}$. Three values of the faint-end LF-slope α are considered. **Panel c:** Contours of $F_{\rm lens}$ as a function of z and $(M_\star - M_{\rm lim})$, assuming[1] $\alpha = -2$.



Supplementary Figure 3. **Probabilities for multiple imaging of $z \simeq 8-10$ galaxy candidates. Panel a:** The probability that a galaxy with observed $M_{AB,1}$ is multiply-imaged. The expected mean number of lenses ($\langle N_{\rm lens}\rangle$) among the $z \simeq 8.6$ and $z \simeq 10.6$ candidates is listed. **Panel b:** The probability that a *lensed* galaxy with observed $M_{AB,1}$ has a corresponding second image with $M_{AB,2} < M_{\rm lim}$. **Panel c:** The fraction of galaxies that are part of a lensed pair in which both images are detectable (68% errors here were computed using a bootstrap method). The expected mean number of systems that would be observed as doubles ($\langle N_{\rm dbl}\rangle$) is listed. We have assumed the determinations of $M_\star = -17.8$ and $\alpha = -2$, and observed absolute magnitudes $M_{AB,1}$ from Yan et al.[4] (the squares correspond to probabilities for the individual galaxy candidates).



Supplementary Figure 4. **Probability distributions for angular proximity and redshift of bright foreground galaxies among the sample of $z \sim 8.6$ candidates. Panel a:** The cumulative distribution for the angular separation between $z \simeq 8.6$ candidates and their nearest foreground galaxies with $H \leq 25$ in the *HUDF* (red histogram). Also shown are the model cumulative distributions of angular separations between *random* lines-of-sight and the nearest bright foreground galaxies (dotted black line), and of angular separations for the brighter image of gravitationally *lensed* objects at $z = 8.6$ (dashed black line). The thick black line shows the composite cumulative distribution generated by summing the *random* and *lensed* histograms, with a weight equal to a lens fraction of $F_{\rm lens} = 0.2$. **Panel b:** The binned histograms (area normalised to unity) for the angular separations of observed candidates (red), for separations in the composite model (thick black), and for separations from random lines of sight (dotted black). **Panel c:** The cumulative redshift distribution for foreground galaxies associated with $z \simeq 8.6$ candidates (red histogram). Also shown are the cumulative distributions for the redshifts of foreground galaxies nearest to *random* lines of sight (dotted black line), and for the expected gravitational lens redshifts assuming sources at $z = 8.6$ (dashed black line). The thick black line shows the composite cumulative distribution ($F_{\rm lens} = 0.2$). We also plot the redshift distribution of foreground galaxies within 1.5 arcseconds of a $z \simeq 8.6$ candidate (blue histogram). **Panel d:** The binned histograms for the foreground galaxy redshifts along lines of sight to dropout candidates (red), and for the composite model (black). In each case, values of $P_{\rm KS}$ corresponding to the comparison of the data with the model distributions are listed.



Supplementary Figure 5. **Probability distributions for angular proximity and redshift of bright foreground galaxies among the sample of $z \sim 10.6$ candidates.** The panels mirror those of Supplementary Figure 4. We assume $F_{\rm lens} = 0.4$ for the model composite distribution.